

Geuvadis - Analysis of alternative splicing

Jean Monlong, Pedro Ferreira, Roderic Guigo

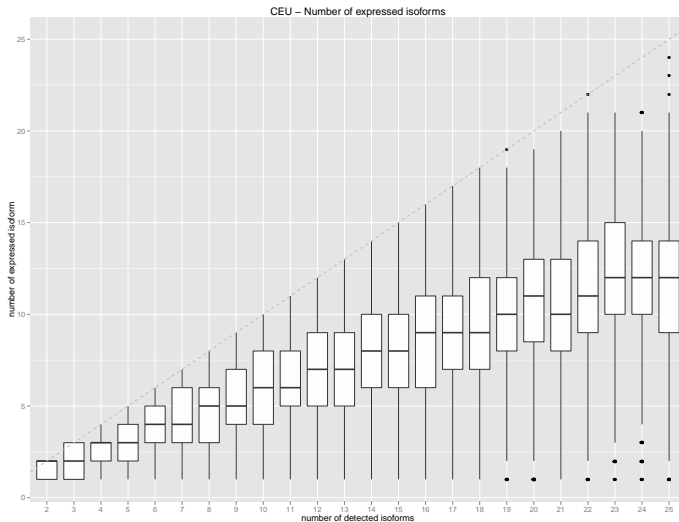
Centre for Genomic Regulation

Barcelona meeting - 9,10th July 2012

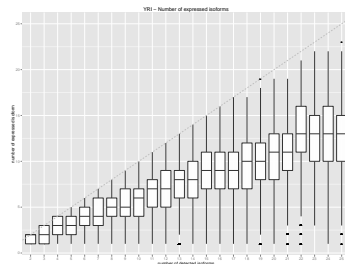
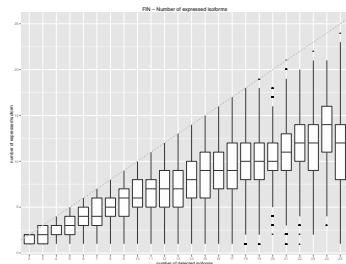
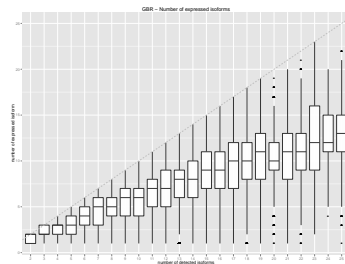
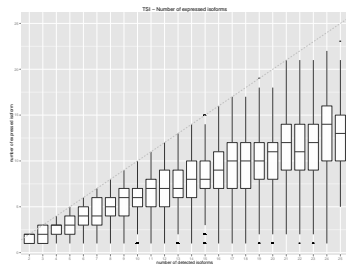
- General splicing statistics.
- Contribution of alternative splicing in the transcript abundance variability.
- Comparison of splicing variability/ratios across the populations.
- Splicing QTL discovery.

General splicing statistics

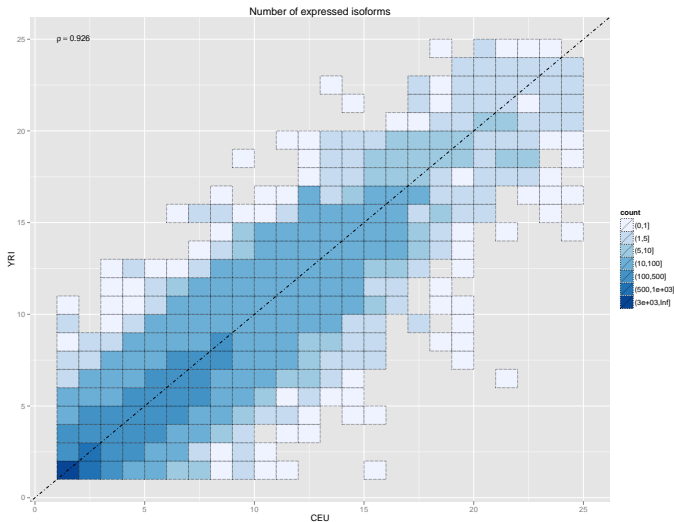
Number of expressed isoforms

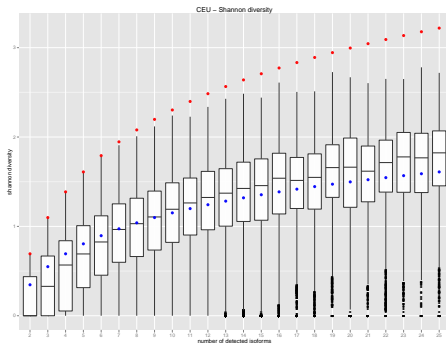
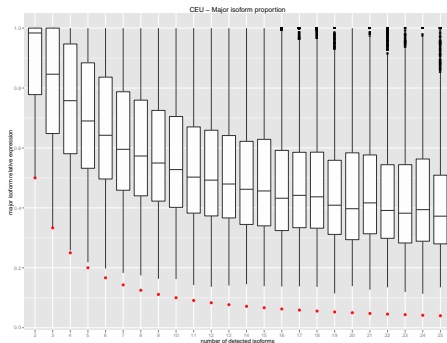


Number of expressed isoforms



Total number of expressed isoforms





Conclusions

- Similar general distribution.
- Some differences at gene level.

Contribution of alternative splicing in the transcript abundance variability.

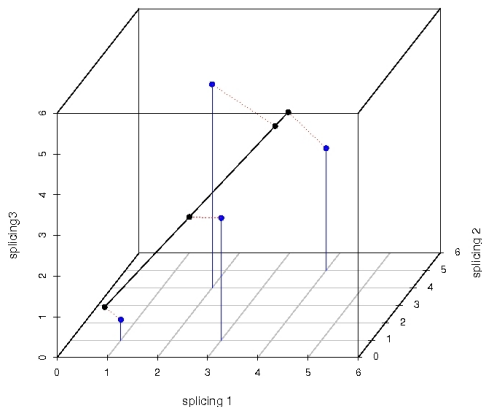
Variability: Gene expression vs Alternative Splicing

Sample A		⇒	Sample B		← Alternative splicing
Isoform 1	10		Isoform 1	90	
Isoform 2	90		Isoform 2	10	
			Sample C		← Gene expression
			Isoform 1	90	
			Isoform 2	810	

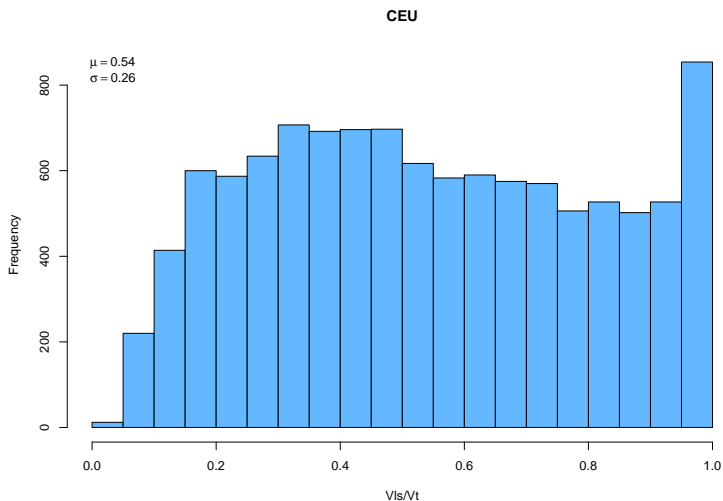
- What is the contribution of gene expression and alternative splicing in the transcript abundance ?

Contribution of alternative splicing in the transcripts' abundance variability

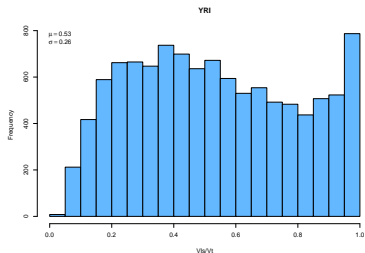
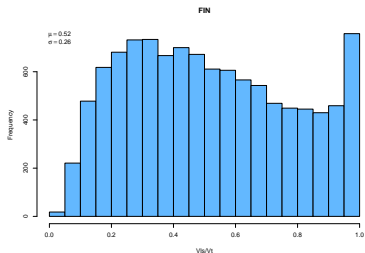
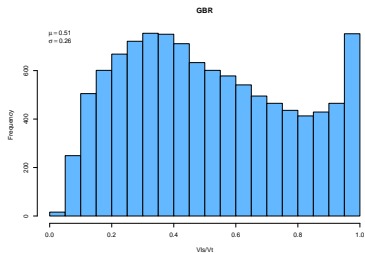
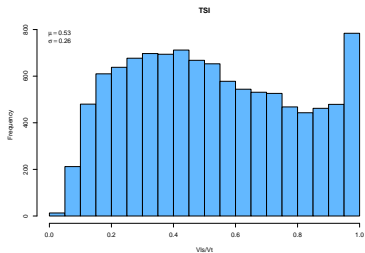
Estimate the ratio $\frac{V_{Is}}{V_t}$ where V_{Is} is the variance when projecting the data to the model of constant splicing ratios and V_t the total variance.



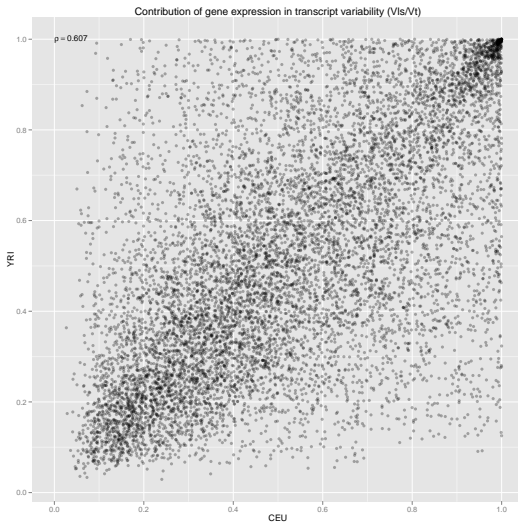
Contribution of gene expression



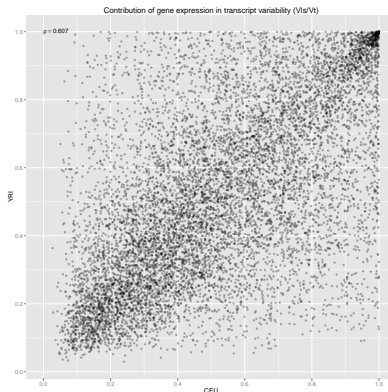
Contribution of gene expression



CEU vs YRI



CEU vs YRI



Conclusions

- More or less similar general distribution.
- Differences at gene level.

Comparison of splicing dispersion/ratios across populations

Method

Analysis steps

- Representation of the samples in the space of the transcript expression ratios.
- Hellinger distance computation for each pair of points.
- MANOVA-like test.
- Multiple-test correction.

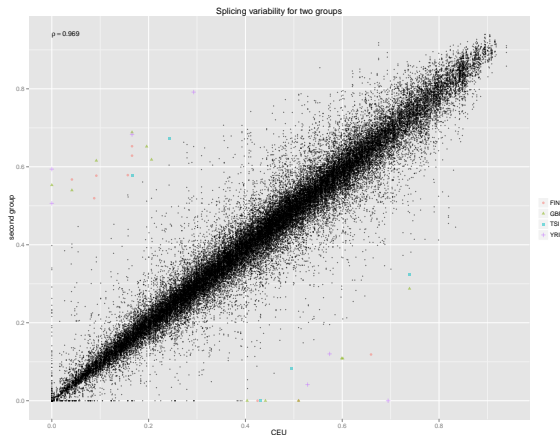
Dispersion comparison then ratios comparison

- Pairwise comparison of the dispersion.
- For each gene, comparison of the splicing ratios across the populations with similar splicing dispersion.

Dispersion comparison: Results

	YRI	CEU	GBR	FIN	TSI
$p_v < 0.01$	263	180	98	70	42
$FDR < 0.01$	73	52	23	12	5

Number of genes with population specific splicing dispersion



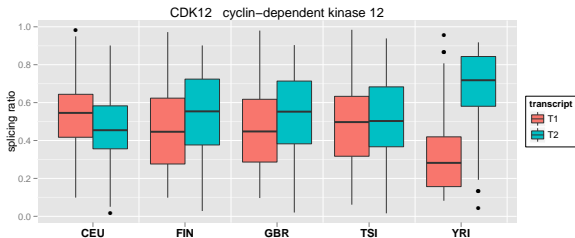
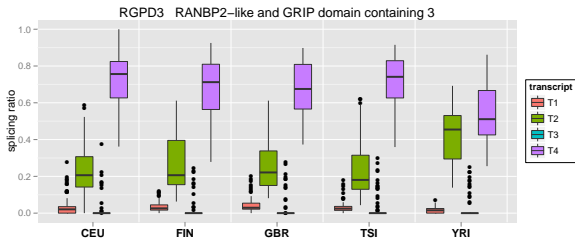
Ratios comparison: Results

	CEU	FIN	GBR	TSI	YRI
$pv < 0.01$	134	92	42	28	537
$FDR < 0.01$	18	19	8	3	257

Number of genes with population specific splicing ratios

-
- Very few population specific splicing ratios found.
 - The more population in the analysis the less population specific.

Ratios comparison: Results - YRI specific



Splicing QTL

- Find SNPs changing the splicing behavior of a gene.
- ANOVA-like method.
- Same as before but comparing the groups created by the samples genotype.

Method

For each gene:

- Computation of a F-ratio like score for each SNPs to study.
- Permutations and computation of the pvalues.
- False discovery control.

Results

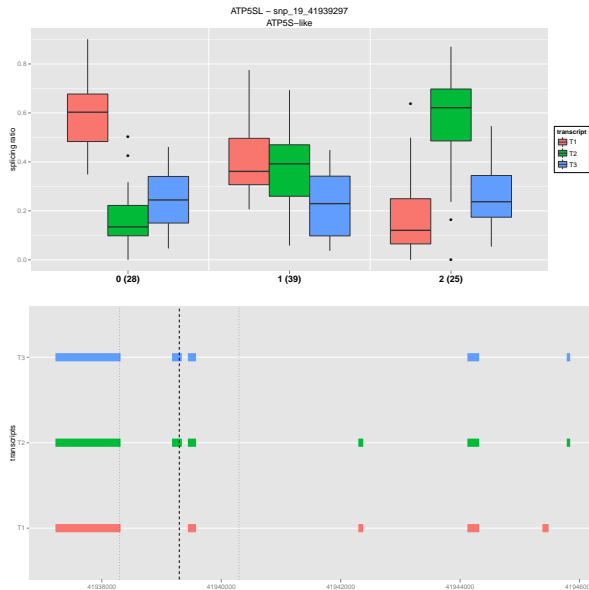
	CEU	FIN	TSI	YRI
Associations (gene-SNP)	5768	3975	1489	8034
Genes	345	186	66	569

Number of significant associations or genes with significant associations after FDR control at 1%.

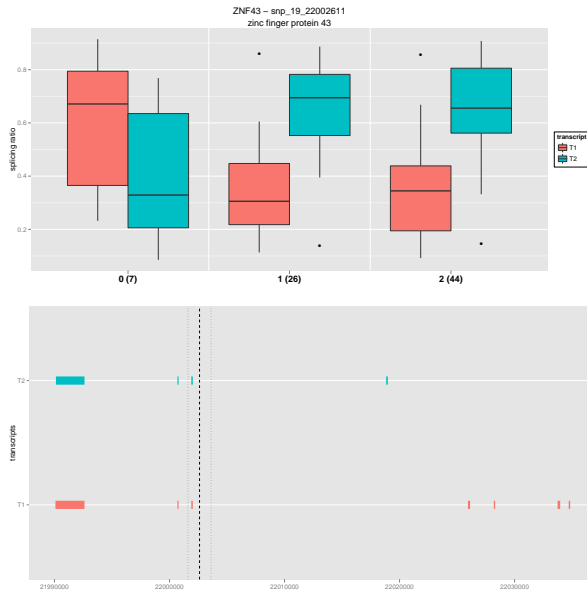
	Exonic	ESS/ESE
CEU	10.5%/8.2%/9.1%	75.6%/75.3%/75.5%
FIN	15.6%/9.1%/9.8%	77%/75.4%/75.5%
TSI	31.1%/10.5%/10.8%	77.2%/75.4%/75.2%
YRI	9%/8.1%/8.6%	76%/75.2%/75.5%

Proportion of exonic SNPs or SNPs into ESS/ESE motives. In green: SNPs with significant association; in black: SNPs with "good" association ($p_v < 0.01$); in red: SNPs with "bad" associations set ($p_v \in [.95, .96]$).

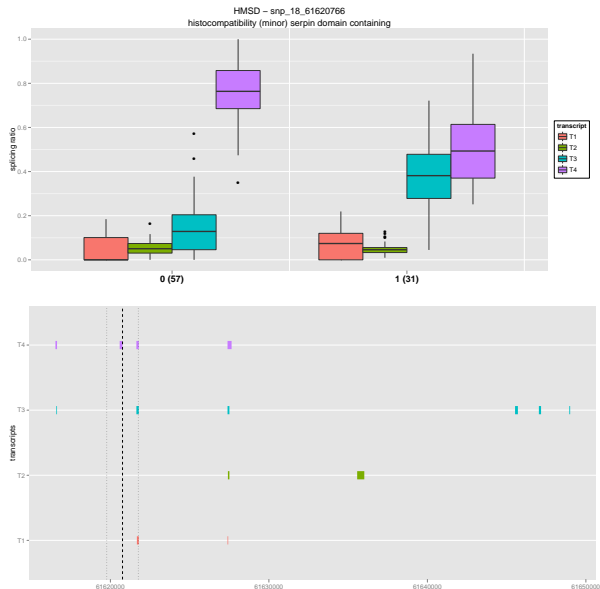
Examples - TSI



Examples - YRI



Examples - TSI

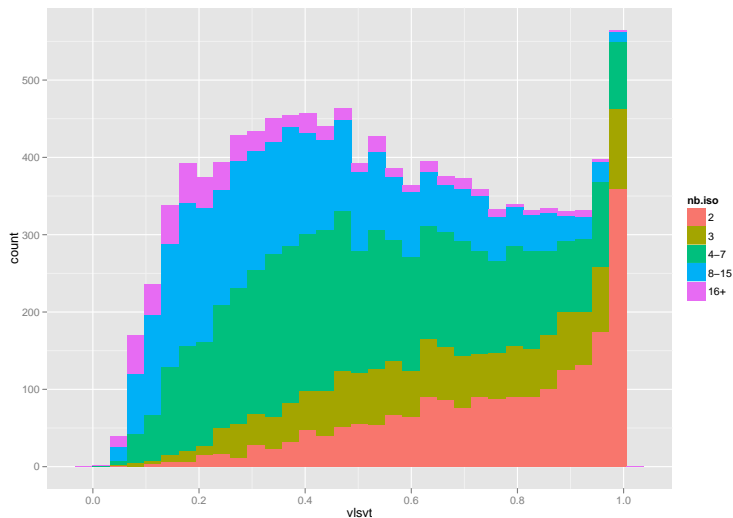


TODO

- Look at the genes with different splicing statistics.
- Deepest analysis of the population comparison splicing dispersion/ratios results..
- Splicing QTL shared by several populations, population specific sQTL.
- ???

Thank you.

Contribution of gene expression



Dispersion comparison: Results

